Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 402 (2005) 74–100
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

On set functions that can be extended to convex functionals

H. Narayanan

Department of Electrical Engineering, Indian Institute of Technology—Bombay, Powai, Mumbai 400076, India

Received 16 February 2004; accepted 15 December 2004
Available online 29 January 2005
Submitted by R.A. Brualdi

Abstract

A set function $f : 2^S \to \Re$, is said to be *polyhedrally tight* (*pt*) (*dually polyhedrally tight* (*dpt*)) iff in the set polyhedron (dual set polyhedron) denoted by P_f (P^f) defined by

$$x(X) \leqslant f(X) \quad \forall X \subseteq S$$
$$(x(X) \geqslant f(X) \quad \forall X \subseteq S),$$

every inequality can be satisfied as an equality (not necessarily simultaneously). We show that these are precisely the set functions that can be extended to convex (concave) functionals over $\Re^S_+$. We characterize such functions and show that if they have certain additional desirable properties, they are forced to become submodular/supermodular. We study *pt* and *dpt* functions using the notion of a *legal dual generator* (*LDG*) structure which is a refinement of the sets of generator vectors of the dual cones associated with the faces of the set polyhedron. We extend $f(g)$ to convex and concave functionals on $\Re^S$ by

$$f_{\text{cup}}(c) \equiv \max_{x \in P_f} c^T x, \quad g_{\text{cap}}(c) \equiv \min_{x \in P^g} c^T x.$$

We then show a refinement (in terms of LDG) of the following discrete separation theorem.

Theorem 0.1. *If f is polyhedrally tight, g is dually polyhedrally tight and $f \geqslant g$ and P_f and P_g have the same dual cones associated with their faces, then $f_{\text{cup}} \geqslant g_{\text{cap}}$ and there exists a modular function h s.t. $f \geqslant h \geqslant g$.*

E-mail address: hn@ee.iitb.ac.in

doi:10.1016/j.laa.2004.12.008

We give sufficient conditions on the dual generator structures of f, g in order that h is integral when f, g are integral. Using these we derive the (integral) Sandwich Theorem for submodular/supermodular functions and (working with a $(0, 1, -1)$ coefficient matrix generalization of set polyhedra), the 1/2-integral Sandwich Theorem for pseudomatroids. We also study the relative positions of Edmonds Intersection Theorem and Frank's Sandwich Theorem in this class of set functions. It turns out that the former is difficult to generalize unless we generalize the definition of convolution while the latter is routinely generalizable to all pt/dpt functions. Using polyhedral ideas we show that if a set function satisfies the Sandwich Theorem with all supermodular functions it must be submodular.

AMS classification: 90C27; 90C46

Keywords: Discrete convexity; Submodular functions; Hahn–Banach Separation Theorem

1. Introduction

In this paper we explore the properties of set functions which can be extended to convex functionals. For this class of general functions we show that the extension can be carried out in a manner similar to the Lovasz extension [5] for submodular functions. These set functions we call polyhedrally tight (pt). Formally, a set function $f : 2^S \to \Re$ is *pt* (dually *pt*) *iff* in the set polyhedron (dual set polyhedron) denoted by P_f (P^f) defined by

$$\begin{array}{ll} x(X) \leqslant f(X) & \forall X \subseteq S \\ (x(X) \geqslant f(X) & \forall X \subseteq S), \end{array}$$

every inequality can be satisfied as an equality (not necessarily simultaneously). We characterize such functions and show that if they have certain additional desirable properties, they are forced to become submodular/supermodular. We study *pt* and *dpt* functions using the notion of a *legal dual generator* (*LDG*) structure which is a refinement of the sets of generator vectors of the dual cones associated with the faces of the set polyhedron.

Our key result is a discrete separation theorem analogous to the Sandwich Theorem [3] for submodular functions. This result states that if f, g are *pt* and *dpt*, have essentially the same LDG structure and $f \geqslant g$ then we can find a modular function h such that $f \geqslant h \geqslant g$. We also prove an 'integrality' version of this theorem provided the associated LDG structure satisfies a much more stringent condition which we call 'hereditary regularity'. The integral Sandwich Theorem for submodular functions is a consequence. By working with $(0, 1, -1)$ vectors a 1/2 integral Sandwich Theorem is derived for pseudo matroids.

Murota and his coworkers have done a careful study of 'discrete convexity' through a series of papers comprehensively described in [7]. Our point of view differs from theirs in that the basic class for us is that of 'polyhedrally tight set' functions or equivalently, the class of set functions that can be extended to convex functionals. The classes that are of interest to them (L- and M-convex functions) are always convex extensible. But the extensions are not necessarily homogeneous and therefore are not necessarily convex functionals. However, when L- and M-convex functions can be extended to convex functionals, they form a strict subclass of polyhedrally tight set functions. Our stand is partially justified by the existence of a separation theorem appropriate for such functions. But it is not clear from our results that the integral separation theorem is valid for a class larger than that of submodular functions.

The outline of the paper is as follows:

Section 2 is on preliminary definitions.

Section 3 defines legal dual generator structures.

Section 4 equates the class of set functions which can be extended to convex functionals to that of polyhedrally tight set functions. It also presents characteristic properties of polyhedrally and dually polyhedrally tight set functions and also some results on what additional properties force such functions to become submodular, supermodular respectively.

Section 5 describes the extension of polyhedrally and dually polyhedrally tight set functions to convex and concave functionals respectively through the use of LDG structures. This is analogous to the Lovasz extension of submodular/supermodular functions.

Section 6 gives the discrete separation theorem and its integral version for polyhedrally and dually polyhedrally tight set functions using compatible LDG structures for the two classes of functions. Also presented is a result which states that the separation theorem fails when such compatibility does not hold. A consequence is that if a class of polyhedrally tight set functions satisfies the separation theorem with all supermodular functions (or Edmond's intersection theorem with all submodular functions) then the class has to be submodular. We also present a generalization of Edmond's intersection theorem for all polyhedrally tight set functions, by generalizing the notion of convolution to the one used in convex analysis.

Section 7 is on conclusions.

2. Preliminaries

Let S be a finite set. We denote the collection of all subsets of S by 2^S and the collection of pairs (X, Y) where X, Y are disjoint subsets of S by 3^S. A function $f : 2^S \to \Re$ is called a *set function* and a function $f : 3^S \to \Re$ is called a *pseudo set function*.

By a vector on S over $\Re$ we mean a mapping $\mathbf{f}$ of S into $\Re$. The *support* of $\mathbf{f}$ is the subset of S over which it takes nonzero values. The *sum* of two vectors $\mathbf{f}, \mathbf{g}$ on S over $\Re$ is defined by $(\mathbf{f} + \mathbf{g})(e_i) \equiv \mathbf{f}(e_i) + \mathbf{g}(e_i)\ \forall e_i \in S$. The *scalar product* of $\mathbf{f}$ by a 'scalar' $\lambda \in \Re$ is a vector $\lambda\mathbf{f}$ defined by $(\lambda\mathbf{f})(e_i) \equiv \lambda(\mathbf{f}(e_i))\ \forall e_i \in S$. The *dot product* of two vectors $\mathbf{f}, \mathbf{g}$ on S over $\Re$ denoted by $\langle \mathbf{f}, \mathbf{g} \rangle$ is defined by

$\langle \mathbf{f}, \mathbf{g} \rangle \equiv \sum_{e \in S} \mathbf{f}(e) \cdot \mathbf{g}(e)$. In place of $\langle \mathbf{f}, \mathbf{g} \rangle$, we would often write $\mathbf{f}^T \mathbf{g}$, treating $\mathbf{f}, \mathbf{g}$ as 'column vectors'. A *cone* is a collection of vectors closed under addition and nonnegative scalar multiplication. It is easily verified that the solution set of $\mathbf{Ax} \leqslant \mathbf{0}$ is a cone. Such cones are said to be *polyhedral*. It is also easily seen that the collection of all vectors which are nonnegative linear combinations of (equivalently *generated by*) a finite set of vectors over $\Re$ forms a cone. Such cones are said to be *finitely generated*. The cone generated by a set of vectors V over $\Re$ is denoted by $C(V)$. We say vectors $\mathbf{x}, \mathbf{y}$ (on the same set S) are *polar* iff $\langle \mathbf{x}, \mathbf{y} \rangle$ (i.e., the dot product) is nonpositive. If $\mathscr{K}$ is a collection of vectors the *polar of* $\mathscr{K}$, denoted by $\mathscr{K}^p$, is the collection of vectors polar to every vector in $\mathscr{K}$. It is easily verified that the polar of a cone is also a cone. Farkas Lemma states:

"Let $\mathscr{C}$ be the polyhedral cone defined by $\mathbf{Ax} \leqslant \mathbf{0}$. A vector $\mathbf{d}$ belongs to $\mathscr{C}^p$ iff the 'row vector' $\mathbf{d}^T$ is a nonnegative linear combination of the rows of $\mathbf{A}$." Equivalently, 'if $\mathscr{C}$ is a finitely generated cone then $\mathscr{C}^{pp} = \mathscr{C}$.' Whenever $\mathscr{C}^{pp} = \mathscr{C}$, we say that $\mathscr{C}$, $\mathscr{C}^p$ are complementary polar.

Let x be a vector in $\Re^S$ (equivalently, on S over $\Re$). Then,

$x(X) \equiv \sum_{e \in X} x(e)$, $X \subseteq S$. The *restriction* of x to $T \subseteq S$ is in $\Re^T$, denoted by x/T and defined by $x/T(e) = x(e)$, $e \in T$. The *restriction* of a set V of vectors in $\Re^S$ to $T \subseteq S$ is denoted by V/T and is the set of restrictions of vectors in V to T. For $X \subseteq S$, χ_X in $\Re^S$ denotes the characteristic vector of X defined by

$$\begin{aligned} \chi_X(e) &= 1, \quad e \in X, \\ &= 0, \quad \text{otherwise} \end{aligned}$$

and for $X, Y \subseteq S$, $X \cap Y = \emptyset$, $\chi_{X,Y}$ denotes the vector defined by

$$\begin{aligned} \chi_{X,Y}(e) &= 1, \quad e \in X, \\ &= -1, \quad e \in Y, \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

The collection of all vectors which are non-negative linear combinations of vectors in V is denoted by $C(V)$. A function $f : \Re^S \to \Re$ is said to be a *convex* (*concave*) *functional* iff

$$\begin{aligned} &f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2), \quad 0 \leqslant \lambda \leqslant 1 \\ &(f(\lambda x_1 + (1-\lambda)x_2) \geqslant \lambda f(x_1) + (1-\lambda) f(x_2), \quad 0 \leqslant \lambda \leqslant 1) \end{aligned}$$

and further $f(\lambda x) = \lambda f(x)$, $\lambda \geqslant 0$. Thus, a convex (concave) functional satisfies

$$\begin{aligned} &f(x_1 + x_2) \leqslant f(x_1) + f(x_2) \\ &(f(x_1 + x_2) \geqslant f(x_1) + f(x_2)). \end{aligned}$$

For a set function $f : 2^S \to \Re$, the polyhedron

$$\begin{aligned} &x(X) \leqslant f(X) \quad \forall X \subseteq S \\ &(x(X) \geqslant f(X) \quad \forall X \subseteq S) \end{aligned}$$

is called the *set polyhedron* (*dual set polyhedron*) associated with $f(\cdot)$. Every line in such polyhedra is bounded in the positive (negative) direction for each variable.

It follows that for every non-negative c in $\Re^S$, $c^T x$ in P_f (P^f) is maximized (minimized) at a vertex.

For a pseudo set function $f : 3^S \to \Re$, the polyhedron

$$\begin{array}{ll} x(X) - x(Y) \leqslant f(X, Y) & \forall X, Y \subseteq S,\ X \cap Y = \emptyset \\ (x(X) - x(Y) \geqslant f(X, Y) & \forall X, Y \subseteq S,\ X \cap Y = \emptyset) \end{array}$$

is called the *pseudo set polyhedron* (*dual pseudo set polyhedron*) associated with $f(\cdot, \cdot)$. Every line in such polyhedra is bounded and therefore the polyhedron is actually a polytope. Thus for any c in $\Re^S$, $c^T x$ in P_f (P^f) is maximized (minimized) at a vertex.

For a set function $f : 2^S \to \Re$, the *restriction* $f \cdot T : 2^T \to \Re$, $T \subseteq S$ is defined by

$$(f \cdot T)(X) = f(X), \quad X \subseteq T,$$

the *contraction* $f \times T : 2^T \to \Re$, $T \subseteq S$ is defined by

$$(f \times T)(X) = f(X \cup (S - T)) - f(S - T), \quad X \subseteq T,$$

the *dual* f^* of f is defined by

$$f^*(X) = f(S) - f(S - X), \quad X \subseteq S.$$

Let $\Pi \equiv S_1, S_2, \ldots, S_k$, be a partition of S. The *fusion* of f relative to Π, denoted by $f_{\text{fus}\Pi}(\cdot)$ is defined on subsets of Π by

$$f_{\text{fus}\Pi}(X_f) \equiv f\left(\bigcup_{T \in X_f} T\right), \quad X_f \subseteq \Pi.$$

A set function $f : 2^S \to \Re$ is said to be *submodular* (*supermodular*) iff

$$\begin{array}{ll} f(X) + f(Y) \geqslant f(X \cup Y) + f(X \cap Y), & \forall X, Y \subseteq S \\ (f(X) + f(Y) \leqslant f(X \cup Y) + f(X \cap Y) & \forall X, Y \subseteq S). \end{array}$$

A pseudo set function $f : 3^S \to \Re$ is said to be a *pseudo matroid* (*pseudo supermodular*) function iff

$$\begin{aligned} f(X_1, Y_1) + f(X_2, Y_2) \geqslant{} & f(X_1 \cup X_2 - Y_1 \cup Y_2, Y_1 \cup Y_2 - X_1 \cup X_2) \\ & + f(X_1 \cap X_2, Y_1 \cap Y_2) \quad X_i, Y_i \subseteq S,\ X_i \cap Y_i = \emptyset \end{aligned}$$

$$\begin{aligned} (f(X_1, Y_1) + f(X_2, Y_2) \leqslant{} & f(X_1 \cup X_2 - Y_1 \cup Y_2, Y_1 \cup Y_2 - X_1 \cup X_2) \\ & + f(X_1 \cap X_2, Y_1 \cap Y_2)) \quad X_i, Y_i \subseteq S,\ X_i \cap Y_i = \emptyset. \end{aligned}$$

A *modular* function $h : 2^S \to \Re$ is both submodular and supermodular. A modular function h such that $h(\emptyset) = 0$ may be identified with a vector h on S, $h(X)$ being the same as $\sum_{e \in X} h(e)$.

3. Legal dual generator (LDG) structures

In this section we introduce the notion of a legal dual generator structure. This is motivated by the structure of the sets of generator vectors of the dual cones at the faces of a polyhedron. These play a key role in extending polyhedrally tight functions to convex functionals analogous to the Lovasz extension [5] of submodular functions. In [8] LDG structures have been called 'shapes' and studied extensively. In the present work, the central idea is a separation theorem. The main overlap of the present work with [8] is in the definition of LDG structures. The notion of extension that is used here is essentially present also in [5], and in [8].

A *legal* (0, 1) *dual generator structure* (*LDG*) $\mathscr{G}$ on S is a collection of sets V of (0, 1) vectors on S, such that

1. If $c \in \Re^S$ and $c \geqslant 0$, then $\exists V \in \mathscr{G}$ and $\lambda_i \geqslant 0$, such that $\sum_i \lambda_i v_i = c$, $v_i \in V$.
2. (Intersection property)—If $V^1, V^2 \in \mathscr{G}$, then $C(V^1 \cap V^2) = C(V^1) \cap C(V^2)$.

For notational convenience, we would often treat a member $V \in \mathscr{G}$ as a matrix whose rows, repetitions not being permitted, are the vectors in V. We define a $(0, 1, -1)$ *LDG structure* $\mathscr{G}$ *on* S similar to the above except that the vectors in V are $(0, 1, -1)$ vectors, and (1) is replaced by (1′).

1′. If $c \in \Re^S$, then $\exists V \in \mathscr{G}$ and $\lambda_i \geqslant 0$, such that $\sum_i \lambda_i v_i = c$, $v_i \in V$.

When members V_i of $\mathscr{G}$ are made up of linearly independent vectors, condition (2) implies that the vectors that generate a given c in different $V_i \in \mathscr{G}$ have to be the same.

We say an LDG structure $\mathscr{G}$ is *hereditary*, if the family $\mathscr{G}/T$ of sets V/T, $V \in \mathscr{G}$ is an LDG structure for every T that is a subset of the underlying set S of $\mathscr{G}$. Given two LDGs $\mathscr{G}_1, \mathscr{G}_2$, we say $\mathscr{G}_1 \geqslant \mathscr{G}_2$ iff for every $V_2 \in \mathscr{G}_2$, there exists a $V_1 \in \mathscr{G}_1$ s.t. $C(V_2) \subseteq C(V_1)$.

Consider the set polyhedron $P_f(P^f)$ defined by

$$
\begin{aligned}
&(\chi_X)^T x \leqslant f(X), \qquad X \subseteq S\\
&((\chi_X)^T x \geqslant f(X), \qquad X \subseteq S).
\end{aligned}
$$

A face F of this polyhedron is defined by imposing the additional condition that some of these inequalities be satisfied as equalities. We associate the corresponding set of row vectors $(\chi_{X_i})^T$ with F and denote it by V_F.

The set of all V_F, where F is a vertex of $P_f(P^f)$ is seen to be a legal dual generator structure which will be denoted by $\mathscr{G}_f$. Another such structure is obtained by taking all faces (instead of only vertices). Similar examples are obtained using the (0,1,−1) set polyhedron. In this case again, $\mathscr{G}_f$ will denote the set of all V_F, where F is a vertex of $P_f(P^f)$.

We say an LDG structure $\mathscr{G}$ is *compatible* with f, iff $\mathscr{G} \leqslant \mathscr{G}_f$. We say an LDG structure $\mathscr{G}$ on S is *regular* iff each $V \in \mathscr{G}$ consists of $|S|$ linearly independent vectors (in particular, the cone $C(V)$ has dimension $|S|$ and therefore nonzero volume in $\Re^{|S|}$). We say $\mathscr{G}$ is *hereditary regular* iff it is regular and the family $\mathscr{G}/T$ of sets V/T, $V \in \mathscr{G}$ is a regular LDG structure for every T that is a subset of the underlying set S of $\mathscr{G}$.

Remark. It may be noted in the case of hereditary regularity that if V is regarded as a matrix of row vectors, the submatrix of V with all rows but columns corresponding to T may contain repeated rows. The *set of rows* of this submatrix would however be linearly independent.

4. Polyhedrally tight set functions

In this section we show that polyhedrally tight (dually polyhedrally tight) set functions are identical to the class of set functions which can be extended to convex (concave) functionals. We show that a natural additional constraint forces such functions to be submodular. We close with a few simple results on *pt* functions derived from a given *pt* function.

Given a set function $f : 2^S \to \Re$, we can naturally associate with it the cone $\mathscr{C}_f$ in $\Re^{|S|+1}$ generated by nonnegative linear combinations of the vectors $(\chi_X, f(X))$, $X \subseteq S$. It is not always possible to recover f from $\mathscr{C}_f$. We show in this section that for the classes of set functions which are extendable to convex (concave) functionals, this is indeed possible through the rule

$$f(X) = \min_{(\chi_X, z) \in \mathscr{C}_f} z \quad \left(f(X) = \max_{(\chi_X, z) \in \mathscr{C}_f} z \right).$$

For such set functions the extension to the convex (concave) functional $\hat{f}$ on $\Re^S_+$ (i.e., on the subset of nonnegative vectors in $\Re^S$) would be by the rule

$$\hat{f}(x) = \min_{(x, z) \in \mathscr{C}_f} z \quad \left(\hat{f}(x) = \max_{(x, z) \in \mathscr{C}_f} z \right).$$

We characterize these set functions by using $\mathscr{C}_f$ and finally through a duality result prove that they are identical to the class of polyhedrally tight (dually polyhedrally tight) set functions.

Theorem 4.1. *Let $f : 2^S \to \Re$ be a set function. Let $\hat{f}$ be defined over $\Re^S_+$ by $\hat{f}(x) \equiv \min_{(x,z) \in \mathscr{C}_f} z$ $(\hat{f}(x) \equiv \max_{(x,z) \in \mathscr{C}_f} z)$. Then*

(a) *$\hat{f}$ is a convex (concave) functional.*
(b) $\hat{f}(y) = \max_{x \in P_f} \langle y, x \rangle$ $(\hat{f}(x) = \min_{x \in P^f} \langle y, x \rangle)$.

Proof. We will consider only the case where $\hat{f}(x) \equiv \min_{(x,z)\in\mathscr{C}_f} z$.

(a) It is clear from the definition of $\mathscr{C}_f$ and that of $\hat{f}$ that $\hat{f}(\lambda x) = \lambda\hat{f}(x), \lambda \geqslant 0$. We therefore need only to show that $\hat{f}(x_1 + x_2) \leqslant \hat{f}(x_1) + \hat{f}(x_2), x_1, x_2 \in \Re^S_+$. We have that $(x_1, \hat{f}(x_1)), (x_2, \hat{f}(x_2)) \in \mathscr{C}_f$ whenever $x_1, x_2 \in \Re^S_+$. But then $(x_1 + x_2, \hat{f}(x_1) + \hat{f}(x_2)) \in C_f$. The desired inequality now follows from the definition of $\hat{f}$.

(b) The cone $\mathscr{C}_f$ is generated by the vectors $(\chi_X, f(X)), X \subseteq S$ through nonnegative linear combinations. The cone $\mathscr{C}^p_f$ is therefore the collection of all (x, α) s.t. $\langle(\chi_X, f(X)), (x, \alpha)\rangle \leqslant 0, X \subseteq S$. Now if $(x, 0) \in \mathscr{C}^p_f$, $\langle\chi_X, x\rangle \leqslant 0, X \subseteq S$ and therefore $x \leqslant 0$. $\mathscr{C}^p_f$ is therefore generated by the following: the vectors $(x, -1)$ which satisfy $\langle(\chi_X, f(X)), (x, -1)\rangle \leqslant 0, X \subseteq S$ (i.e., the vectors $(x, -1)$, $x \in P_f$), the vectors $(x, +1)$ which satisfy $\langle(\chi_X, f(X)), (x, +1)\rangle \leqslant 0, X \subseteq S$ and finally, nonnegative linear combination of the 'negative directions' $(-\chi_{\{e_i\}}, 0)$, $e_i \in S$.

Let $g : \Re^S_+ \to \Re$ be defined by $g(y) \equiv \max_{x\in P_f}\langle y, x\rangle$. Now when $y \geqslant 0$, $(y, \hat{f}(y)) \in \mathscr{C}_f$. So $\langle(y, \hat{f}(y)), (x, -1)\rangle \leqslant 0, x \in P_f$, i.e., $\langle y, x\rangle \leqslant \hat{f}(y), x \in P_f$. Hence $\hat{f}(y) \geqslant \max_{x\in P_f}\langle y, x\rangle = g(y)$.

By the definition of $g(y)$, $\langle(y, g(y)), (x, -1)\rangle \leqslant 0$, whenever $x \in P_f$, i.e., whenever $(x, -1) \in \mathscr{C}^p_f$. Further, since $y \geqslant 0$ we have $\langle(y, g(y)), (x, 0)\rangle \leqslant 0$ whenever $x \leqslant 0$, i.e., whenever $(x, 0) \in \mathscr{C}^p_f$. Next we have

$$\begin{aligned}
g(y) &\equiv \max_{x\in P_f}\langle y, x\rangle \\
&= \max\langle y, x\rangle, \quad \text{for } \langle\chi_X, x\rangle \leqslant f(X),\ \forall X \subseteq S \\
&= \min\sum\lambda_i f(X_i), \quad \text{for } \sum\lambda_i\chi_{X_i} = y,\ \lambda_i \geqslant 0 \\
&\leqslant \max\ \sum\lambda_i f(X_i), \quad \text{for } \sum-\lambda_i\chi_{X_i} = -y,\ \lambda_i \geqslant 0 \\
&= \min\langle -y, \tilde{x}\rangle, \quad \text{for } \langle-\chi_X, \tilde{x}\rangle \geqslant f(X),\ \forall X \subseteq S \\
&= \min\langle -y, \tilde{x}\rangle, \quad \text{for } \langle(\chi_X, f(X)), (\tilde{x}, 1)\rangle \leqslant 0,\ \forall X \subseteq S \\
&= \min\langle -y, \tilde{x}\rangle, \quad \text{for } (\tilde{x}, 1) \in \mathscr{C}^p_f \\
&= -\max\langle y, \tilde{x}\rangle, \quad \text{for } (\tilde{x}, 1) \in \mathscr{C}^p_f
\end{aligned}$$

(where we have used LP duality twice, noting that in each case the dual program has an optimum). Thus $\langle((y, g(y)), (\tilde{x}, 1)\rangle \leqslant 0$, for $(\tilde{x}, 1) \in \mathscr{C}^p_f$. We have thus shown that $\langle((y, g(y)), (x, \alpha)\rangle \leqslant 0$, whenever (x, α) is in a specified set of generators for $\mathscr{C}^p_f$. Hence $\langle((y, g(y)), (x, \alpha)\rangle \leqslant 0$, for $(x, \alpha) \in \mathscr{C}^p_f$, i.e., $(y, g(y)) \in \mathscr{C}^{pp}_f$. By Farkas Lemma, since $\mathscr{C}_f$ is finitely generated, $\mathscr{C}^{pp}_f = \mathscr{C}_f$. So $(y, g(y)) \in \mathscr{C}_f$. Hence, by the definition of $\hat{f}(y)$, $\hat{f}(y) \leqslant g(y)$. Thus $\hat{f}(y) = g(y)$ as required. □

Remark. We have captured P_f in the above proof as the set of all vectors $(x, -1)$ in C^p_f. To handle P^f it is convenient to consider the set of all vectors $(x, -1)$ in $-\mathscr{C}^p_f$.

Theorem 4.2. *Let $f : 2^S \to \Re$ be a set function. Then* (a) *f can be extended to a convex* (*concave*) *functional $\tilde{f} : \Re^S_+ \to \Re$ iff for every $X \subseteq S$, and any $\lambda_i \geqslant 0$ s.t. $\sum \lambda_i \chi_{X_i} = \chi_X$, we have, $\sum \lambda_i f(X_i) \geqslant f(X)(\sum \lambda_i f(X_i) \leqslant f(X))$.*

(b) *If f can be extended to a convex* (*concave*) *functional $\tilde{f}$, then $\hat{f}(x) \equiv \min_{(x,z)\in\mathscr{C}_f} z$ ($\hat{f}(x) \equiv \max_{(x,z)\in\mathscr{C}_f} z$), is a convex* (*concave*) *functional extension of f and $\tilde{f} \leqslant \hat{f}$ ($\tilde{f} \geqslant \hat{f}$).*

Proof. We will only consider the convex case.

(a) (Only if) We have, $\tilde{f}(\chi_X) = f(X)$, $X \subseteq S$. Since $\tilde{f}$ is a convex functional, $\sum \lambda_i f(X_i) = \sum \lambda_i \tilde{f}(\chi_{X_i}) \geqslant \tilde{f}\left(\sum \lambda_i \chi_{X_i}\right) = f(X)$, whenever $\lambda_i \geqslant 0$ and are s.t. $\sum \lambda_i \chi_{X_i} = \chi_X$.

(If) Let, for every $X \subseteq S$, $\sum \lambda_i f(X_i) \geqslant f(X)$ whenever $\lambda_i \geqslant 0$ and are s.t. $\sum \lambda_i \chi_{X_i} = \chi_X$. Define $\hat{f}(x) \equiv \min_{(x,z)\in\mathscr{C}_f} z$. We will verify that $\hat{f}$ is an extension of f and that it is a convex functional. Clearly $(\chi_X, \hat{f}(\chi_X)) \in \mathscr{C}_f$. Now by the definition of $\mathscr{C}_f$, there exist χ_{X_i}, $i = 1, \ldots, k$, λ_i, $i = 1, \ldots, k$, $\lambda_i \geqslant 0$, $\sum \lambda_i \chi_{X_i} = \chi_X$ and $\sum \lambda_i f(X_i) = \hat{f}(\chi_X)$. Therefore $\hat{f}(\chi_X) \geqslant f(X)$. On the other hand by the definition of $\hat{f}$, $\hat{f}(\chi_X) \leqslant f(X)$. Hence $\hat{f}(\chi_X) = f(X)$. Thus $\hat{f}$ is an extension of f.

Next, to prove the convexity of $\hat{f}$, let $x_1, x_2 \in \Re^S_+$ and let $0 \leqslant \lambda \leqslant 1$. We have $(x_1, \hat{f}(X_1)), (x_2, \hat{f}(X_2)) \in \mathscr{C}_f$. Hence $((\lambda x_1 + (1-\lambda)x_2, \lambda \hat{f}(x_1) + (1-\lambda) \hat{f}(x_2)) \in \mathscr{C}_f$. Therefore, by the definition of $\hat{f}$, we must have $\lambda \hat{f}(x_1) + (1-\lambda) \hat{f}(x_2) \geqslant \hat{f}((\lambda x_1) + (1-\lambda)x_2)$, which proves that $\hat{f}$ is a convex functional.

(b) That $\hat{f}$ is a convex functional extension of f follows immediately from the '(if)' part of the proof above. Suppose for some $x \in \Re^S$, $\tilde{f}(x) > \hat{f}(x)$. By the definition of $\hat{f}$, it follows that $(x, \hat{f}(x)) \in \mathscr{C}(f)$. But then it is possible to find χ_{X_i}, $i = 1, \ldots, k$, λ_i, $i = 1, \ldots, k$, $\lambda_i \geqslant 0$, $\sum \lambda_i \chi_{X_i} = x$ and $\sum \lambda_i f(X_i) = \sum \lambda_i \tilde{f}(\chi_{X_i})$ $= \hat{f}(x) < \tilde{f}\left(\sum \lambda_i (\chi_{X_i})\right) = \tilde{f}(x)$. This contradicts the fact that $\tilde{f}$ is a convex functional. □

We can now equate the classes of *pt* functions and of set functions extendable to convex functionals.

Theorem 4.3. *A set function $f : 2^S \to \Re$ is pt* (*dpt*)

(a) *iff it can be extended to a convex* (*concave*) *functional over $\Re^S_+$.*
(b) *iff*

$$\sum \lambda_i f(X_i) \geqslant f(X) \qquad X_i \subseteq X \subseteq S$$
$$(\sum \lambda_i f(X_i) \leqslant f(X) \qquad X_i \subseteq X \subseteq S),$$

whenever $\lambda_i \geqslant 0$ *and are s.t.* $\sum \lambda_i \chi_{X_i} = \chi_X$.

Proof. We will consider only the *pt* case. By definition,
f is *pt* iff $f(X) = max_{x \in P_f} \langle \chi_X, x \rangle, \forall X \subseteq S$,
i.e., iff $f(X) = \hat{f}(\chi_X) \equiv \min_{(\chi_X, z) \in \mathscr{C}_f} z, \forall X \subseteq S$ (by Theorem 4.1),
i.e., iff f has $\hat{f}$ as an extension,
i.e., iff f has a convex functional extension (by Theorem 4.2(b)).
(b) follows from (a) above and Theorem 4.2(a). □

The next result describes an operation on a set function that is analogous to the Dilworth Truncation operation on a submodular function.

Theorem 4.4. *Let* $f : 2^S \to \Re$ *be a set function. Let* $f_*(X) \equiv \min_{\sum \lambda_i \chi_{X_i} = \chi_X, \lambda_i \geqslant 0} \sum \lambda_i f(X_i)$. *Then* (a) $P_f = P_{f_*}$. (b) f_* *is pt.* (c) *If* $f \geqslant f'$ *and* f' *is pt then* $f_* \geqslant f'$.

Proof. (a) It is easily seen that $\mathscr{C}_f \supseteq \mathscr{C}_{f_*}$, equivalently, $\mathscr{C}_f^p \subseteq \mathscr{C}_{f_*}^p$. Hence (since $P_f \equiv \{x, (x, -1) \in C_f^p\}$), $P_f \subseteq P_{f_*}$. However in the '$\leqslant$' inequalities defining P_f, P_{f_*} the right side for any inequality for P_f is always greater or equal to the right side of the corresponding inequality for P_{f_*}. Hence $P_f \supseteq P_{f_*}$. It follows that $P_f = P_{f_*}$.

(b) $\hat{f}_*(x) \equiv \min_{(x,z) \in \mathscr{C}_{f_*}} z, x \in \Re_+^S$ is a convex functional by Theorem 4.1. It is clear, by the definition of f_*, that it is an extension of f_*. Therefore by Theorem 4.3, f_* is pt.

(c) If f' is pt, by Theorem 4.3, it can be extended to a convex functional. So by Theorem 4.2, $f'(X) \leqslant \min_{\sum \lambda_i \chi_{X_i} = \chi_X, \lambda_i \geqslant 0} \sum \lambda_i f'(X_i) \leqslant \min_{\sum \lambda_i \chi_{X_i} = \chi_X, \lambda_i \geqslant 0} \sum \lambda_i f(X_i) = f_*(X)$. □

We next show among other things that polyhedrally tight functions, if they satisfy certain additional properties, are forced to become submodular functions. We need the following standard duality result on linear inequalities which can be proved using Farkas Lemma.

Theorem 4.5. *The system of inequalities*

$$A_1 X \leqslant b_1,$$
$$A_2 X = b_2$$

has a solution iff, whenever

$$\lambda_1^T A_1 + \lambda_2^T A_2 = 0, \quad \lambda_1 \geqslant 0,$$

we also have,

$$\lambda_1^T b_1 + \lambda_2^T b_2 \geqslant 0.$$

Theorem 4.6. *Let f be a set function on subsets of S.*

(a) *If f is pt, then, every restriction of f is pt and every fusion of f is pt.*
(b) *f is submodular if the dual of every restriction of f is dpt.*
(c) *f is submodular if every contraction of f is pt.*

Proof. (a) is immediate from the definition of restriction and fusion. (b) Let $T \subseteq S$. Then

$$\begin{aligned}(f \cdot T)^*(X) &= (f \cdot T)(T) - (f \cdot T)(T - X), \quad X \subseteq T\\ &= f(T) - f(T - X), \quad X \subseteq T.\end{aligned}$$

From Theorem 4.3, $(f \cdot T)^*$ is *dpt* iff

$$\sum \lambda_i (f \cdot T)^*(X_i) \leqslant (f \cdot T)^*(X), \quad X \subseteq T,\ X_i \subseteq X,\ \forall i,$$

whenever $\lambda_i \geqslant 0$ and are s.t., $\sum \lambda_i \chi_{X_i} = \chi_X$. In particular, choose X_1, X_2 s.t.

$$X_1 \cup X_2 = X \subseteq T, \quad X_1 \cap X_2 = \emptyset.$$

We then have

$$(f \cdot T)^*(X_1) + (f \cdot T)^*(X_2) \leqslant (f \cdot T)^*(X),$$

i.e.,

$$f(T) - f(T - X_1) + f(T) - f(T - X_2) \leqslant f(T) - f(T - X),$$

i.e.,

$$f(T - X_1) + f(T - X_2) \geqslant f(T) + f(T - X).$$

Thus, for any $A, B \subseteq S$ if we choose

$$T = A \cup B, \quad X = T - (A \cap B), \quad X_1 = T - A, \quad X_2 = T - B,$$

it would follow that $X_1 \cup X_2 = X$ and $X_1 \cap X_2 = \emptyset$ and

$$f(A) + f(B) \geqslant f(A \cup B) + f(A \cap B).$$

The result follows.

(c) We have

$$(f \times T)(X) = -f(S - T) + f((S - T) \cup X), \quad X \subseteq T \subseteq S.$$

We choose X_1, X_2 s.t.

$$X_1 \cup X_2 = X \subseteq T, \quad X_1 \cap X_2 = \emptyset.$$

Since $f \times T$ is *pt*, by Theorem 4.3, we have,

$$(f \times T)(X_1) + (f \times T)(X_2) \geqslant (f \times T)(X),$$

i.e.,

$$-f(S-T) + f((S-T) \cup X_1) - f(S-T) + f((S-T) \cup X_2) \\ \geqslant -f(S-T) + f((S-T) \cup X),$$

i.e.,

$$f((S-T) \cup X_1) + f((S-T) \cup X_2) \\ \geqslant f(S-T) + f((S-T) \cup X) \quad X_1, X_2 \subseteq X \subseteq T.$$

Thus for any $A, B \subseteq S$, if we choose $S - T = A \cap B$, $(S-T) \cup X = A \cup B$, $X_1 = A \cap T$, $X_2 = B \cap T$, we have

$$f(A) + f(B) \geqslant f(A \cap B) + f(A \cup B). \qquad \square$$

We remark that a submodular function which takes zero value on the null set is *pt*. Also every restriction and contraction of submodular functions which take zero value on the null set are also submodular and take zero value on the null set. So parts (b) and (c) of Theorem 4.6 yield characterizations of submodular functions which take zero value on the null set. The following elementary result is about some situations where the *pt* property is inherited by derived functions.

Theorem 4.7. *If the set function f is pt on subsets of S. Then,*

(a) *the set function g, where*

$$g(X) \equiv f(X), \quad X \subseteq S, \quad g(X) \equiv \min(k, f(S)), \quad X = S, \textit{ is pt}.$$

(b) *Let, $c \in \Re^S$, $c > 0$. Let*
$f_c(X) \equiv \max_{x \in P_f} c_X^T x$, *where*, $c_X(e) = c(e)$ $e \in X$, $\equiv 0$ *otherwise. Then* $f_c(\cdot)$ *is pt.*
(c) $x + f$, *where x is a vector in $\Re^S$ is pt. Also* $\mathcal{G}_f = \mathcal{G}_{f+x}$.
(d) *The set function f is pt iff $-f$ is dpt. Further,* $\mathcal{G}_f = \mathcal{G}_{-f}$.
(e) *If f_1, f_2 are pt then $f_1 + f_2$ is pt.*

Proof. (a) follows immediately from Theorem 4.3.

(b) Consider the polyhedron

$$x(X) \leqslant f_c(X), \quad X \subseteq S.$$

Now $f_c(X) = \max_{x \in P_f} c_X^T x$. Thus there exists a vector $\hat{x} \in P_f$ s.t. $c_X^T \hat{x} = f_c(X)$. Define x_c by

$$x_c(e) = c(e) \cdot \hat{x}(e) \quad e \in S.$$

Clearly $x_c(X) = f_c(X)$ and

$$x_c(Y) = \sum_{e \in Y} c(e)\hat{x}(e) \leqslant \max_{x \in P_f} c_Y^T x = f_c(Y).$$

Thus f_c is pt.

(c) We have $\max_{y \in P_{f+x}} c^T y = c^T x + \max_{y \in P_f} c^T y$ and $c^T y$ maximizes at $\hat{x}$ in P_f iff $c^T(y+x)$ maximizes at $\hat{x}+x$ in P_{f+x}. The additional result follows by the definition of $\mathscr{G}_f, \mathscr{G}_{f+x}$.

(d) The result follows from the fact that $c^T x$ maximizes at $\hat{x}$ in P_f iff $c^T(-x)$ minimizes at $-\hat{x}$ in P^{-f}.

(e) is routine. □

5. Extension of *pt* and *dpt* functions

Polyhedrally tight and dually polyhedrally tight set functions that can be extended to convex and concave functionals can be described either 'primally' as in Theorem 4.2 using the cone $\mathscr{C}_f$ or 'dually' using $\mathscr{C}_f^p$ or equivalently using P_f, P^f as described below. By Theorem 4.1 the resulting extensions would be the same. We will henceforth adopt the 'dual' approach. Instead of using $\hat{f}$ to denote the convex/concave extension of f, we will henceforth use f_{cup} for the convex extension of a pt function and f_{cap} for the concave extension of a dpt function.

Let $f: 2^S \to \Re$ be a pt (dpt) function. We extend f to $f_{\text{cup}}: \Re^S \to \Re$ ($f_{\text{cap}}: \Re^S \to \Re$) by

$$f_{\text{cup}}(c) \equiv \max_{x \in P_f} c^T x,$$
$$f_{\text{cap}}(c) \equiv \min_{x \in P^f} c^T x.$$

For vectors c, for which the above maximum (minimum) does not exist, we take

$$f_{\text{cup}}(c) = +\infty \quad (f_{\text{cap}}(c) = -\infty).$$

One can similarly define f_{cup} (f_{cap}) in the case where $f: 3^S \to \Re$ is ppt ($dppt$). When f is pt, since

$$\left(\max_{x \in P_f} (\lambda c_1^T + (1-\lambda)c_2^T)x\right) \leqslant \max_{x \in P_f} \lambda c_1^T x + \max_{x \in P_f} (1-\lambda) c_2^T x, \quad 0 \leqslant \lambda \leqslant 1,$$

it is clear that f_{cup} is convex. Since the inequality would be reversed in the case of P^f for a dpt function f, we would have that f_{cap} is concave.

Let f be pt. Let $\hat{x}$ be a vertex of P_f at which $c^T x$ is maximum. Let the coefficient vectors of the supporting hyperplanes at $\hat{x}$ form the rows v_i', $i = 1, \ldots, n$, of $V' \in \mathscr{G}_f$ and let $v_i' \equiv \chi_{Xi}$. we then have

$$V'\hat{x} = \begin{bmatrix} f(X_1) \\ \vdots \\ f(X_n) \end{bmatrix}, \quad c = \lambda^T V', \quad \lambda \geqslant 0.$$

Hence,

$$
\begin{aligned}
f_{\text{cup}}(c) &= c^T \hat{x} = (\lambda^T V')\hat{x} \\
&= \lambda^T \begin{bmatrix} f(X_1) \\ \vdots \\ f(X_n) \end{bmatrix}.
\end{aligned}
$$

The above procedure for computing $f_{\text{cup}}(c)$ will yield a unique output for $c \geqslant 0$, even if c maximizes at more than one vertex of P_f, since $\mathscr{G}_f$ satisfies conditions (1) and (2) of LDG structures. Further, since f is pt, by the definition of $f_{\text{cup}}(c)$ it is clear that

$$
f_{\text{cup}}(\chi_X) = f(X), \quad X \subseteq S.
$$

Now let $\mathscr{G} \leqslant \mathscr{G}_f$ and let $c \in C(V) \in \mathscr{G}$. There exists $V' \in \mathscr{G}_f$ s.t. $C(V) \subseteq C(V')$. Let $V = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$. We then have, by condition (1) of LDG structures

$$
v_j = \sigma_j^T \begin{bmatrix} v_1' \\ \vdots \\ v_n' \end{bmatrix}, \quad \sigma_j \geqslant 0, \quad j = 1, \ldots, k.
$$

Therefore, as seen earlier,

$$
f_{\text{cup}}(v_j) = \sigma_j^T \begin{bmatrix} f_{\text{cup}}(v_1') \\ \vdots \\ f_{\text{cup}}(v_n') \end{bmatrix}.
$$

Now since $c \in C(V) \in \mathscr{G}$, $\exists \mu \geqslant 0$ s.t.

$$
c^T = \mu^T \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix} = \mu^T \begin{bmatrix} \sigma_1^T \\ \vdots \\ \sigma_k^T \end{bmatrix} \begin{bmatrix} v_1' \\ \vdots \\ v_n' \end{bmatrix}.
$$

Thus

$$
f_{\text{cup}}(c) = \mu^T \begin{bmatrix} \sigma_1^T \\ \vdots \\ \sigma_k^T \end{bmatrix} \begin{bmatrix} f_{\text{cup}}(v_1') \\ \vdots \\ f_{\text{cup}}(v_n') \end{bmatrix} = \mu^T \begin{bmatrix} f_{\text{cup}}(v_1) \\ \vdots \\ f_{\text{cup}}(v_k) \end{bmatrix}.
$$

Now, if vectors in all the members of $\mathscr{G}$ are (0, 1) vectors, then each v_i is χ_{Y_i} for some $Y_i \in S$. As noted above,

$$
f_{\text{cup}}(v_i) = f(Y_i).
$$

Thus,

$$f_{\text{cup}}(c) = \mu^T \begin{bmatrix} f(Y_1) \\ \vdots \\ f(Y_k) \end{bmatrix},$$

where $\chi_{Y_i} = v_i$. It follows that the procedure for computing $f_{\text{cup}}(c)$ would yield the same value whether we use $\mathscr{G}$ or $\mathscr{G}_f$ as long as $\mathscr{G} \leqslant \mathscr{G}_f$.

A similar discussion can be carried out for the case of ppt and $dppt$ functions.

We summarize the above discussion in the form of a theorem.

Theorem 5.1. *Let $\mathscr{G}$ be an LDG structure and let* $c = \lambda^T \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $\lambda \geqslant 0$ *for some* $\{v_1, \ldots, v_n\} \in \mathscr{G}$.

(a) *Let f be a pt (dpt) function such that $\mathscr{G}$ is compatible with it. Then,*

$$f_{\text{cup}}(c) = \lambda^T \begin{bmatrix} f(X_1) \\ \vdots \\ f(X_n) \end{bmatrix} \left(f_{\text{cap}}(c) = \lambda^T \begin{bmatrix} f(X_1) \\ \vdots \\ f(X_n) \end{bmatrix} \right),$$

where, $\chi_{X_i} = v_i$, $i = 1, \ldots, n$.

(b) *Let f be a ppt ($dppt$) function such that $\mathscr{G}$ is compatible with it. Then,*

$$f_{\text{cup}}(c) = \lambda^T \begin{bmatrix} f(X_1, Y_1) \\ \vdots \\ f(X_n, Y_n) \end{bmatrix} \left(f_{\text{cap}}(c) = \lambda^T \begin{bmatrix} f(X_1, Y_1) \\ \vdots \\ f(X_n, Y_n) \end{bmatrix} \right),$$

$$X_i \cap Y_i = \emptyset, \quad i = 1, \ldots, n,$$

where, $\chi_{X_i} - \chi_{Y_i} = v_i$, $i = 1, \ldots, n$.

Remark. We thus see that the idea of the Lovasz extension defined for submodular functions [4,5] carries over to pt, ppt functions using compatible LDG structures.

An immediate corollary to the above theorem is the following.

Corollary 5.1. *Let f be a pt (ppt) and g, a dpt ($dppt$) function on subsets of S. Let there exist an LDG structure $\mathscr{G}$ compatible with both f and g. If $f \geqslant g$, then $f_{\text{cup}} \geqslant g_{\text{cap}}$.*

Proof. We consider only the pt, dpt case.

For any $c = \lambda^T \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $\{v_1, \ldots, v_n\} \in \mathscr{G}$, $\lambda \geqslant 0$, we have

$$f_{\text{cup}}(c) = \lambda^T \begin{bmatrix} f(X_1) \\ \vdots \\ f(X_n) \end{bmatrix},$$
$$g_{\text{cap}}(c) = \lambda^T \begin{bmatrix} g(X_1) \\ \vdots \\ g(X_n) \end{bmatrix},$$

since $v_i = \chi_{X_i}, i = 1, \ldots, n$. Since $f(X) \geqslant g(X)\ \forall X \subseteq S$, and $\lambda \geqslant 0$, it follows that $f_{\text{cup}}(c) \geqslant g_{\text{cap}}(c)$. $\square$

Example S. If f is a submodular function on subsets of S, the LDG structure $\mathscr{G}^s$ naturally compatible with it is obtained as follows:

Let V_σ, σ a permutation of $\{1, 2, \ldots, S\}$, be composed of the row vectors

$$\begin{aligned} v_j(\sigma(i)) &= 1, \quad i \leqslant j, \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

Then, $\mathscr{G}^s \equiv \{V_\sigma, \sigma \text{ a permutation of } \{1, 2, \ldots, S\}\}$. It is well known that $\mathscr{G}^s \leqslant \mathscr{G}_f$, and also that if $\mathscr{G}_f \geqslant \mathscr{G}^s$, and f is pt then f is submodular.

Example P [6]. If f is a pseudomatroid function on subsets of S,

$$\begin{aligned} &f(X_1, Y_1) + f(X_2, Y_2) \\ &\quad \geqslant f(X_1 \cup X_2 - Y_1 \cup Y_2, Y_1 \cup Y_2 - X_1 \cup X_2) + f(X_1 \cap X_2, Y_1 \cap Y_2) \\ &\quad\quad \forall X_i, Y_i \subseteq S, X_i \cap Y_i = \emptyset. \end{aligned}$$

We enlarge the above LDG structure $\mathscr{G}^s$ as follows to obtain the LDG $\mathscr{G}^p$, naturally compatible with f. Let $\alpha : \{1, 2, \ldots, S\} \to \{1, -1\}$. Let $V_{\sigma\alpha}$ be composed of the row vectors

$$\begin{aligned} V_j(\sigma(i)) &= \alpha(i) \quad i \leqslant j, \\ &= 0 \quad\quad \text{otherwise.} \end{aligned}$$

Then, $\mathscr{G}^p \equiv \{V_{\sigma\alpha}, \sigma$ a permutation of $\{1, 2, \ldots, S\}, \alpha : \{1, 2, \ldots, S\} \to \{1, -1\}$.

6. The discrete separation Theorem for *pt* functions

In this section we prove a discrete separation theorem, which says that a modular function lies between a pt and a suitably related dpt function. For submodular functions one can show that this result is equivalent to Edmonds' intersection theorem [2] which, in our language, says that the convolution of two pt submodular functions is pt. But the latter result is probably characteristic of submodular functions while the discrete separation theorem is fundamental for all pt functions. We also give sufficient conditions for the existence of an integral modular function between integral pt and dpt functions. Using these, we prove the integral sandwich theorem for submodular functions and the $1/2$-integral sandwich theorem for pseudo matroid functions.

The key idea in our proof of the discrete separation theorem is to extend the pt, dpt functions to convex and concave functions and to use the Hahn–Banach Separation Theorem. This extension can be carried out provided we have a common compatible LDG structure for both functions. To round out the discussion, we prove that the discrete separation theorem does indeed fail when there is no such LDG structure.

For convenience of proof, we combine both the general discrete separation theorem and the integral version in the same theorem. We need the following lemma in the proof of the integral version.

Lemma 6.1. *Let C_1, C_2 be cones in $\Re^S$ with generator sets R_1, R_2 respectively with the vectors in R_1, R_2 having entries in $\{0, 1, -1\}$. Further let*

$$C(R_1/T \cap R_2/T) = C(R_1/T) \cap C(R_2/T), \quad \forall T \subseteq S.$$

Let $d(\cdot,\cdot) : C_1 \times C_2 \to \Re$ be a linear functional. Let $S = Y \uplus Z \uplus W$ where Z is a singleton and W is not void. Further, let every vector in $C(R_1/W \cap R_2/W)$ be expressible uniquely as a nonnegative linear combination of vectors in $(R_1/W \cap R_2/W)$. Then, if the extreme (maximum or minimum) value of $d(\cdot,\cdot)$ among all pairs of vectors (x_1, x_2) s.t.

$$x_1/W = x_2/W$$
$$(x_1 - x_2)/Z = 1$$

exists then it also occurs on pairs (r_1, r_2) where

(a) *if R_1, R_2 are composed of $(0, 1)$ vectors, then $r_1 \in R_1,\ r_2 \in R_2$.*
(b) *if R_1, R_2 are composed of $(0, 1, -1)$ vectors, then*

$$r_1 \text{ or } 2r_1 \in R_1,$$
$$r_2 \text{ or } 2r_2 \in R_2.$$

Proof. We will consider the case where $d(\cdot,\cdot)$ is to be minimized. Maximization case can be handled similarly.

Let $d(\cdot,\cdot)$ reach a minimum at (r_1, r_2) under the conditions $r_1 \in C_1,\ r_2 \in C_2$ $r_1/W = r_2/W, \quad (r_1 - r_2)/Z = 1.$ Since $C(R_1/W \cap R_2/W) = C(R_1/W) \cap C(R_2/W)$, and every vector in $C(R_1/W \cap R_2/W)$ is expressible uniquely as a nonnegative linear combination of vectors in $(R_1/W \cap R_2/W)$, it is possible to build two matrices A, B on column set S, rows members of R_1, R_2 respectively or zero s.t. the sub matrices A^W, B^W of A, B, on column sets W and on all rows, are identical, and further

$$(r_1|r_2) = \lambda^T [A \mid B],$$

where λ has all positive entries. We will show that a pair (r_1, r_2) that minimizes $d(\cdot,\cdot)$ can be chosen so that the matrix $(A \mid B)$ has only one row.

Claim 1. *There need be no row* $(A_i \mid B_i)$ *with* $A_i/Z = B_i/Z = 0$.

Observe that $d(\cdot,\cdot)$ cannot take negative value on such a row as otherwise it can be arbitrarily decreased. Hence, setting λ_i to zero cannot increase $d(\cdot,\cdot)$ while still satisfying $(r_1 - r_2)/Z = 1, r_1/W = r_2/W$. This proves the claim.

We will henceforth assume (A, B) does not have such rows.

Claim 2. *There need be no row* (A_i, B_i) *with* $A_i/Z - B_i/Z < 0$.

Let $\mu_j \equiv (A_j - B_j)/Z \; \forall j$. Further, let $\mu_j > 0,\ j \leqslant k$ and $\mu_j < 0,\ j > k$. Since $(r_1 - r_2)/Z = 1$, we must have $\sum_j \mu_j \lambda_j = 1$. We need to minimize $\sum_j d_j \lambda_j, \lambda_j \geqslant 0$, where $d_j \equiv d(A_j, B_j)$ under the condition $\sum_j \mu_j \lambda_j = 1$. Alternatively, we need to minimize $\sum_j d'_j \lambda'_j$, where $d'_j \equiv d_j/\mu_j, \lambda'_j \equiv \mu_j \lambda_j, \lambda_j \geqslant 0$ under the condition $\sum_j \lambda'_j = 1$. We note that $\lambda'_j \geqslant 0,\ j \leqslant k$ and $\lambda'_j \leqslant 0,\ j > k$. Clearly, in order to minimize $\sum_j d'_j \lambda'_j$, we cannot do better than by taking atmost two of the λ'_js to be non-zero, corresponding to the least value of $d'_j,\ j \leqslant k$ and the greatest value of $d'_j,\ j > k$. We may suppose only $\lambda'_t > 0, t \leqslant k$ and $\lambda'_r < 0, r > k$. We thus need to minimize $d'_t\lambda'_t + d'_r\lambda'_r$ given that $\lambda'_t + \lambda'_r = 1, \lambda'_t \geqslant 0, \lambda'_r \leqslant 0$. This expression has a minimum provided $d'_t\lambda'_t + d'_r\lambda'_r \geqslant 0$ for $\lambda'_t + \lambda'_r = 0, \lambda'_t \geqslant 0, \lambda'_r \leqslant 0$. But then the minimum occurs at $\lambda'_t = 1, \lambda'_r = 0$. This proves that a pair (r_1, r_2) exists that minimizes $d(\cdot,\cdot)$ which, when expressed as a linear combination of rows of A, B does not use any row (A_i, B_i) with $A_i/Z - B_i/Z < 0$. This proves the claim.

Claim 3. *There need be only one row in* $(A|B)$.

When $(A|B)$is a 0, 1 matrix $(r_1|r_2)$ can be taken to be this row. When $(A|B)$ is a 0, 1, −1 matrix $(r_1|r_2)$ or $2(r_1|r_2)$ can be taken to be this row.

By the foregoing, We may take all rows to have $(A_i - B_i)/Z > 0$.

Case 1: Let $(A|B)$ be a 0,1 matrix. In this case $A_i/Z = 1,\ B_i/Z = 0$.
Since $(r_1 - r_2)/Z = 1$, (r_1, r_2) is therefore a convex combination of rows of $(A|B)$. Therefore $d(\cdot,\cdot)$ reaches its minimum at one of the rows of $(A|B)$. This proves the claim for this case.

Case 2: Let $(A|B)$ be a 0, 1, −1 matrix. We would have

$$\begin{array}{ll} A_i/Z = 1, & B_i/Z = 0 \\ A_i/Z = 0, & B_i/Z = -1 \end{array} \quad type1$$

or

$$A_j/Z = 1, \quad B_j/Z = -1 \quad type2$$

$$(r_1|r_2) = \sum_{type1} \lambda_i A_i + \sum_{type2} \lambda_j (A_j/2) \quad \lambda_i, \lambda_j > 0$$

clearly $(r_1|r_2)$ is a convex combination of the $(A_i|B_i)$ and the $(A_j/2, B_j/2)$. Once again, the minimum will occur at one of these rows. This proves the claim for this case.

The lemma is thus proved. □

We now prove the discrete separation theorem. We follow a standard proof of the Hahn–Banach Separation Theorem (see for instance [1]).

Theorem 6.1. *Let $f : 2^S \to \Re$, $g : 2^S \to \Re$ ($f : 3^S \to \Re$, $g : 3^S \to \Re$) be set (pseudoset), pt and dpt (pseudo pt and pseudo dpt) functions respectively. Let $f \geqslant g$ and let there be a regular LDG structure $\mathcal{G}$ compatible with f, g.*

(a) *Then, there exists a modular function h s.t. $f \geqslant h \geqslant g$.*
(b) *If, further, f, g are integral, every vertex of P_f, P^g is integral and $\mathcal{G}$ is a heriditary LDG structure, then h can be chosen to be integral when f, g are set functions and can be chosen to be 1/2-integral when f, g are pseudo set functions.*

Proof. Let $p, q : \Re^S \to \Re$ denote $f_{\text{cup}}, g_{\text{cap}}$ respectively. By Corollary 5.1 $p \geqslant q$. Now, let $Y \subset S$ and let Z be a singleton set contained in $S - Y$. We will denote by y, y' vectors on S whose support is contained in Y and by z the vector defined by

$$z(e) = 1, \quad e \in Z,$$
$$ = 0, \quad \text{otherwise}$$

Let m be a linear functional on the subspace of vectors y on $\Re^S$ whose support $\text{supp}(y)$ is contained in Y. Let m be s.t.

$$m(y) \leqslant p(x+y) - q(x), \quad \forall x \in \Re^S, \ \forall y \in \Re^S, \ \text{ s.t. supp}(y) \subseteq Y. \qquad (*)$$

By induction on the size of Y we will show that there exists a linear functional m on $\Re^S$ s.t.

$$m(y) \leqslant p(x+y) - q(x), \quad \forall x \in \Re^S, \ \forall y \in \Re^S.$$

Clearly when the size of Y is zero, i.e., when $Y = \emptyset$, any linear functional m satisfies the above condition. We will show that m can be extended to a linear functional (also denoted m) on the subspace of vectors y on $\Re^S$ whose support $\text{supp}(y)$ is contained in $Y \uplus Z$. This we do by first showing that the value $m(z)$ of the extended functional at the vector z, defined as above, can be chosen in such a way that

$$m(y+z) \leqslant p(x+y+z) - q(x), \quad \forall x \in \Re^S, \ \forall y \in \Re^S, \ \text{ s.t. supp}(y) \subseteq Y. \qquad (**)$$

$$m(y'-z) \leqslant p(x'+y'-z) - q(x'),$$
$$\forall x' \in \Re^S,\ \ \forall y' \in \Re^S,\ \ \text{s.t. } \mathrm{supp}(y') \subseteq Y. \qquad (***)$$

Now, by our assumption

$$p(x+x'+y+y') - q(x+x') \geqslant m(y+y'),$$
$$\forall (x+x') \in \Re^S,\ \ \forall (y+y') \in \Re^S,\ \ \text{s.t. } \mathrm{supp}(y+y') \subseteq Y.$$

Since p, q are convex and concave functionals

$$\text{LHS} \leqslant p(x+y+z) - q(x) + p(x'+y'-z) - q(x') \quad \text{and}$$
$$\text{RHS} = m(y) + m(y').$$

Hence,

$$p(x+y+z) - q(x) - m(y) \geqslant -p(x'+y'-z) + q(x') + m(y')$$
$$\forall y, y' \in \Re^S,$$

with $\mathrm{supp}(y, y') \subseteq Y,\ \forall x, x' \in \Re^S$. Hence, $m(z)$ can be chosen so that

$$p(x+y+z) - q(x) - m(y) \geqslant m(z) \geqslant -p(x'+y'-z) + q(x') + m(y'), \qquad (\surd)$$

i.e.

$$p(x+y+z) - q(x) - m(y+z) \geqslant 0 \geqslant q(x') - p(x'+y'-z) + m(y'-z)$$

(since the extended m is a linear functional), i.e.,

$$p(x+y+z) - q(x) - m(y+z) \geqslant 0,$$
$$p(x'+y'-z) - q(x') - m(y'-z) \geqslant 0, \qquad (\surd\surd)$$
$$\forall x, x' \in \Re^S,\ \ y, y' \in \Re^S,\ \ \mathrm{supp}(y, y') \subseteq Y.$$

For any vector $\hat{y}$ with support in $Y \uplus Z$, we have $\hat{y} = y \pm tz$ with $t \geqslant 0$. Since

$$p(\lambda x) = \lambda p(x) \quad \lambda \geqslant 0,$$
$$q(\lambda x) = \lambda q(x) \quad \lambda \geqslant 0,$$

$$p(x+y \pm tz) = tp(x/t + y/t \pm z), \quad t > 0,$$
$$q(x) = tq(x/t), \quad t > 0.$$

Hence, $(\surd\surd)$ implies that

$$p(x+\hat{y}) - q(x) - m(\hat{y}) \geqslant 0, \quad \forall x \in \Re^S,\ \hat{y} \in \Re^S,\ \mathrm{supp}(\hat{y}) \subseteq Y \uplus Z.$$

Repeating this process a finite number of times we get,

$$p(x+y) - q(x) - m(y) \geqslant 0, \quad \forall x, y \in \Re^S.$$

Setting $x = 0$, we get $p(y) \geqslant m(y), \forall y \in \Re^S$ and setting $y = -x$, we get $q(x) \leqslant m(x), \forall x \in \Re^S$. Set $h(X) \equiv m(\chi_X)$, if f, g are set functions and $h(X, Y) \equiv m(\chi_X - \chi_Y)$, if f, g are pseudo set functions. We then have $f(X) = p(\chi_X) \geqslant m(\chi_X) = h(X) \geqslant q(\chi_X) = g(X)$, when f, g are set functions and $f(X, Y) = p(\chi_X - \chi_Y) \geqslant m(\chi_X - \chi_Y) = h(X, Y) \geqslant q(\chi_X - \chi_Y) = g(X, Y)$ when f, g are pseudo set functions. This proves (a).

(b) Here, we need to start with an integral (or 1/2-integral in the case of pseudo *pt* function) linear functional and keep enlarging its domain satisfying ($\surd$) so that it remains integral or 1/2-integral as the case might be. When $Y = \emptyset$, since $m(0) = 0$, we start with an integral functional. To enlarge its domain appropriately, essentially as in ($\surd$), if we show that minimum value of LHS and the maximum value of the RHS are integers (or 1/2-integers in the case of pseudo *pt* function), our task is done. We will do this for the minimum value of LHS, the other can be proved similarly.

Let x_2 denote x, x_1 denote $x + y + z$. By definition

$$p(x_1) = \max_{\alpha \in P_f} x_1^T \alpha,$$

$$q(x_2) = \min_{\alpha \in P^g} x_2^T \alpha.$$

Let $p(x_1) - q(x_2) - m(y)$ reach a minimum at $\hat{x_1}, \hat{x_2}, \hat{y}$ under the condition that $(x_1 - x_2)/W = 0, (x_1 - x_2)/Z = 1$. Therefore, $p(\hat{x_1}), q(\hat{x_2})$ are finite. By the hypothesis of part(b) of the theorem, there exist integer vertices say v_f, v_g of P_f, P^g s.t. $p(\hat{x_1}) = {v_f}^T \hat{x_1}$ and $q(\hat{x_2}) = {v_g}^T \hat{x_2}$. Let the dual cones corresponding to v_f, v_g be $\mathscr{C}^f, \mathscr{C}^g$ with $\hat{x_1} \in \mathscr{C}^f$ and $\hat{x_2} \in \mathscr{C}^g$. Since $\mathscr{G}$ is compatible with f, g respectively there exist $V_1, V_2 \in \mathscr{G}$ s.t.

$$\hat{x_1} \in \mathscr{C}(V_1) \subseteq \mathscr{C}^f \quad \text{and} \quad \hat{x_2} \in \mathscr{C}(V_2) \subseteq \mathscr{C}^g.$$

Then the expression

$$({v_f}^T)x_1 - ({v_g}^T)x_2 - m(y)$$

reaches a minimum at $\hat{x_1}, \hat{x_2}, \hat{y}$ under the condition that

$$x_1 \in \mathscr{C}(V_1), \quad x_2 \in \mathscr{C}(V_2), \quad (x_1 - x_2)/W = 0 \quad (x_1 - x_2)/Z = 1.$$

Define $\hat{m}(x_1, x_2) \equiv m(y)$, where $\mathrm{supp}(y) \subseteq Y$, $y/Y = (x_1 - x_2)/Y$. Clearly $\hat{m}$ is a linear functional on x_1, x_2 since m is a linear functional. It follows that $({v_f}^T)x_1 - ({v_g}^T)x_2 - m(y) = ({v_f}^T)x_1 - ({v_g}^T)x_2 - \hat{m}(x_1, x_2)$ is a linear functional, say $d(\cdot, \cdot)$, on (x_1, x_2). The linear functional $d(x_1, x_2)$ reaches a minimum at $(\hat{x_1}, \hat{x_2})$ under the condition that $x_1 \in \mathscr{C}(V_1), \quad x_2 \in \mathscr{C}(V_2), \quad (x_1 - x_2)/W = 0 \quad (x_1 - x_2)/Z = 1$.

Further, under this condition, if it reaches a minimum at any (x_1', x_2'), since $d(x_1', x_2') = d(\hat{x_1}, \hat{x_2})$, therefore $p(x_1) - q(x_2) - m(y)$ reaches a minimum at (x_1', x_2') under the (less restrictive) condition $(x_1 - x_2)/W = 0, (x_1 - x_2)/Z = 1$.

Now $\mathscr{G}$ is hereditary regular. Hence V_1, V_2 satisfy the hypothesis of Lemma 6.1. Hence by Lemma 6.1, when $\mathscr{G}$ is (0, 1) LDG, $d(x_1, x_2)$ reaches its minimum at some (x_1', x_2'), where x_1', x_2' are vectors in V_1, V_2 and when $\mathscr{G}$ is a (0, 1, −1) LDG, reaches its minimum at some (x_1', x_2') where either x_1', x_2' or $2x_1', 2x_2'$ are vectors in V_1, V_2.

Since v_f, v_g are integers, it follows that the minimum value of $p(x_1) - q(x_2) - m(y) = (v_f)^T x_1 - (v_g)^T x_2 - m(y)$, where $(x_1 - x_2)/W = 0, (x_1 - x_2)/Z = 1$, $\text{supp}(y) \subseteq Y$, $y/Y = (x_1 - x_2)/Y$, is integer in the case where f, g are set functions and 1/2- integer, in the case where f, g are pseudo set functions.

By a similar proof the same statement can be made about the RHS of $(\surd)$. It follows that $m(z)$ can be chosen to be an integer in the case of set functions and 1/2 integer in the case of pseudo set functions. Thus the set Y can be enlarged to $Y \uplus Z$ with the property that

$$m(y) \leqslant p(x+y) - q(x), \quad \forall x, y \in \Re^S, \ \text{supp}(y) \subseteq Y \uplus Z$$

m integral in the case of set functions and 1/2 integral in the case of pseudo set functions. Repeating this process a finite number of times yields the required result. □

We observe that the (0, 1) LDG structure $\mathscr{G}^s$ (Example S) compatible with submodular and supermodular functions $f, g : 2^S \rightarrow \Re$ is hereditary regular. So if $f \geqslant g$ and f, g are integral, there is an integral modular function h s.t.

$$f \geqslant h \geqslant g.$$

We also note that the $(0, 1, -1)$ LDG structure (Example P) compatible with pseudo matroid and pseudo supermodular function $f, g : 3^S \rightarrow \Re$ is also hereditary regular. So if $f \geqslant g$ and f, g are integral, there is a 1/2 integral modular function h s.t.

$$f \geqslant h \geqslant g.$$

The next result says that the Discrete Separation Theorem can be stated in an equivalent but different form. These equivalent forms would therefore hold whenever the Discrete Separation Theorem holds, i.e., whenever there exists an LDG compatible with those associated with the relevant pt/dpt functions. Here, the first part of the result is concerned with equivalence of the conditions in (a), (b). By Theorem 6.1 we know that (b) is true when there exists a $\mathscr{G}$ such that $\mathscr{G} \leqslant \mathscr{G}_i, i = 1, 2$. It follows from the result that (a) is also true when there exists a $\mathscr{G}$ such that $\mathscr{G} \leqslant \mathscr{G}_i, i = 1, 2$. This latter result is due to Sohoni [8]. However, neither separation nor integrality results are considered by that author.

Theorem 6.2. *Let f_1, f_2 be pt (ppt) on subsets of S. Then the following are equivalent*:

(a) $P_{f_1+f_2} = P_{f_1} + P_{f_2}$
(b) *There exists a modular function h s.t. $f_1 \geqslant h \geqslant x - f_2$, whenever $f_1 \geqslant x - f_2$ for some modular function x.*

Further, if (a), (b) *hold the following are equivalent*:

(a′) f_1, f_2 *are integral and every integral vector in* $P_{f_1+f_2}$ *is the sum of an integral vector in* P_{f_1} *and an integral vector in* P_{f_2}.
(b′) *If* f_1, f_2, x *are integral and* $f_1 \geqslant x - f_2$, *then there exists an integral* h *s.t.* $f_1 \geqslant h \geqslant x - f_2$.

Proof. We consider only the *pt* case, the *ppt* case being similar. Let (a) hold. Suppose $f_1 \geqslant x - f_2$. Then $0 \in P_{f_1+f_2-x}$. So there exists an x' s.t $x' \in P_{f_1}$ and $-x' \in P_{f_2-x}$. Clearly we have $f_1 \geqslant x' \geqslant x - f_2$. If, further, (a′) holds, since 0 is integral x' can be chosen to be integral so that (b′) holds.

Next, it is clear that $P_{f_1+f_2} \supseteq P_{f_1} + P_{f_2}$ whether or not (b) holds. Let (b) hold.

Let $x \in P_{f_1+f_2}$. We have $x(X) \leqslant (f_1 + f_2)(X), \forall X \subseteq S$. Consider the *dpt* function $g = x - f_1$. Now $f_2 \geqslant x - f_1 = g$. Hence, there exists a vector h s.t.

$$f_2(X) \geqslant h(X) \geqslant g(X), \quad \forall X \subseteq S.$$

Hence,

$$x - h \leqslant x - g = f_1$$

and $h \leqslant f_2$.

It follows that $x \in P_{f_1} + P_{f_2}$. Thus (a) holds.

Further, let (b′) hold. If f_1, f_2, x are integral, $f_2, x - f_1$ are integral, so h can be taken to be integral. Thus $x = (x - h) + h$, the RHS vectors being integral vectors in P_{f_1}, P_{f_2} respectively. Thus (a′) holds. □

The next theorem appears to indicate that common compatible LDG has to exist for pt, dpt functions in order that the Discrete Separation Theorem holds.

Theorem 6.3. *Let* f, g *be pt, dpt (ppt, dppt) functions respectively on subsets of* S. *Let* $\mathscr{G}_f, \mathscr{G}_g$ *be LDGs but let* $\mathscr{G}_f \not\geqslant \mathscr{G}_g$ *and let* $\mathscr{G}_g$ *be regular. Then there exists a modular function* α *s.t.* $f \geqslant g + \alpha$, *but such that there is no modular function between* f *and* $g + \alpha$.

Proof. We will only consider the pt, dpt case. We first observe that $f' \geqslant h \geqslant g'$ where f' is *pt* and g' is *dpt*, iff the vector h on S, where $\sum_{e \in X} h(e) = h(X)$, is in $P_{f'} \cap P^{g'}$. We will build the required modular function α s.t. $f \geqslant g + \alpha$ but $P_f \cap P^{g+\alpha} = \emptyset$. Since $\mathscr{G}_f \not\geqslant \mathscr{G}_g$ there exists $V' \in \mathscr{G}_g$ not contained in any $V \in \mathscr{G}_f$. Since $\mathscr{G}_g$ is regular (i.e. every member made up of $|S|$ independent vectors) and since an LDG has the property that every nonnegative vector is in some cone $\mathscr{C}(V)$, where V belongs to the LDG, and there are only a finite number of such cones, there exists $V \in \mathscr{G}_f$ s.t. $\mathscr{C}(V) \cap \mathscr{C}(V')$ has nonzero volume. Now, $V \neq V'$ and V' has $|S|$

vectors. Hence, $|V \cap V'| < |S|$. So $\mathscr{C}(V \cap V')$ has zero volume and is not equal to $\mathscr{C}(V) \cap \mathscr{C}(V')$.

By the definition of $\mathscr{G}_f$, $\mathscr{G}_g$, there exist vertices v_f, v_g whose supporting hyperplanes $\chi_{Xi}^T x = f(X_i)$, $\chi_{Xj}^T x = g(X_j)$ are s.t. the χ_{Xi} make up the vectors in V and the χ_{Xj} make up the vectors in V'.

Add a modular function β, which takes zero value on the null set, to g s.t. $v_f = v_g + \beta$ (β denoting the vector corresponding to the modular function β). Let us denote $g + \beta$ by g'.

Now, by Theorem 4.7(c) $\mathscr{G}_g = \mathscr{G}_{g'}$. So $V' \in \mathscr{G}_{g'}$. Next, since $\mathscr{C}(V) \cap \mathscr{C}(V')$ has nonzero volume, there exists a vector b in the interior of $\mathscr{C}(V) \cap \mathscr{C}(V')$. Such a vector is in the interior of the cones $\mathscr{C}(V), \mathscr{C}(V')$. Clearly, $b^T x,\ x \in P_f$ maximizes only at v_f and $b^T x,\ x \in P^{g'}$ minimizes only at $v_g + \beta = v_f$. Thus, the hyperplane $b^T x = b^T v_f$ touches P_f, $P^{g'}$ only at v_f and separates P_f and $(P^{g'} - v_f)$.

Now since $v_f \in P_f \cap P^{g'}$

$$g'(X) \leqslant \chi_X^T v_f \leqslant f(X) \quad \forall X \subseteq S.$$

Further, $f(X) = g'(X)$ iff $g'(X) = \chi_X^T v_f = f(X)$, i.e. iff $\chi_X \in V \cap V'$. Let $f(X) - g'(X) > \epsilon > 0$ when $\chi_X \notin V \cap V'$. Let λ be a nonzero vector orthogonal to all the vectors in $V \cap V'$. Such a vector exists since $|V \cap V'| \neq |S|$. Choose the direction of λ s.t. $\lambda^T b > 0$ and its magnitude s.t. $|\lambda^T \chi_X| < \epsilon\ \forall X \subseteq S$. Clearly, $f(X) - (g' + \lambda)(X) > 0$ whenever $\chi_X \notin V \cap V'$ and since $\lambda^T \chi_X = 0$, $\chi_X \in V \cap V'$,

$$f(X) = (g' + \lambda)(X), \quad \chi_X \in V \cap V'.$$

Thus, $f \geqslant g' + \lambda$. Consider $P^{g'+\lambda}$. This no longer intersects P_f since v_f which was the only common vertex at which P_f and $P^{g'}$ intersected has split into vertices v_f and $v_f + \lambda$ in P_f, $P^{g'+\lambda}$ respectively. The hyperplane $b^T x = b^T v_f$ now separates P_f and $P^{g'+\lambda}$ with $b^T v_f < b^T(v_f + \lambda)$. Let $\alpha = \beta + \lambda$. We thus have $f \geqslant g + \alpha$ but $P_f \cap P^{g+\alpha} = \emptyset$. Hence, there is no modular function h s.t. $f \geqslant h \geqslant g + \alpha$. $\square$

We now consider some consequences of Theorem 6.3. Suppose f is pt and satisfies the Discrete Separation Theorem 6.1 for every g that is dpt, supermodular and satisfies $g \leqslant f$. It is easy to build a supermodular function for which $g \leqslant f$ and $\mathscr{G}_g = \mathscr{G}^s$ ($\mathscr{G}^s$ as described in Example S). Theorem 6.3 would then imply that $\mathscr{G}_f \geqslant \mathscr{G}^s$. But this implies that f is submodular.

Next let f be *pt* and satisfy Edmonds' intersection theorem [2] with every submodular function f_2 that is *pt*. i.e.

$$\left[\max_{x \in P_f \cap P_{f_2}} x(S) = \min_{X \subseteq S} (f(X) + f_2(S - X)) \right].$$

From this we can infer that f must be submodular as follows.

Suppose g is dpt and supermodular and $f \geqslant g$. It is easy to show that $g^*(X)$ $\left[= g(S) - g(S - X)\right]$ is submodular and further $g^*(\emptyset) = 0$. Hence g^* is *pt* and submodular.

By Edmonds intersection theorem

$$\begin{aligned}\max\left\{x(S)|x \in P_f \cap P_{g^*}\right\} &= \min_{X\subseteq S}\left\{f(X)+g^*(S-X)\right\}\\ &= \min_{X\subseteq S}\{f(X)+g(S)-g(X)\}\\ &= g(S).\end{aligned}$$

Consider the vector $x \in P_f \cap P_{g^*}$ s.t. $x(S) = g(S)$. We have

$$x(X) \leqslant f(X) \quad \forall X \subseteq S.$$

Further,

$$\begin{aligned}x(S-X) &\leqslant g^*(S-X) \quad \forall X \subseteq S\\ &\leqslant g(S)-g(X) \quad \forall X \subseteq S\\ &\leqslant x(S)-g(X) \quad \forall X \subseteq S,\end{aligned}$$

i.e.,

$$x(X) \geqslant g(X) \quad \forall X \subseteq S.$$

Thus, $f \geqslant x \geqslant g$. This means that f satisfies the discrete separation theorem with every supermodular *pt* function g and is therefore submodular.

Next we consider whether Edmonds' Intersection Theorem can be generalized to pt functions. If one relaxes the definition of convolution we have the following result which does not even need the two pt functions to be compatible in terms of the associated LDG structures. But with the standard definition of convolution ($f_1 * f_2(X) \equiv \min_{Y\subseteq X}(f_1(Y) + f_2(X-Y))$) the generalization even without integrality conditions appears difficult for any class other than submodular functions.

Theorem 6.4. *Let f_1, f_2 be pt. Then $f_1 * f_2$ is pt, where*

$$f_1 * f_2(X) \equiv \min_{c_1+c_2=\chi_X}((f_1)_{\mathrm{cup}}(c_1) + (f_2)_{\mathrm{cup}}(c_2)).$$

Proof. $\mathbf{P_{f_1*f_2} \subseteq P_{f_1} \cap P_{f_2}}$
Clearly, $P_{f_1} \cap P_{f_2}$ is defined by the inequalities $\chi_X^T x \leqslant \min(f_1(X), f_2(X)) \geqslant \min_{c_1+c_2=\chi_X}(f_1\,\mathrm{cup}(c_1) + f_2\,\mathrm{cup}(c_2))$. It is thus clear that for each inequality of $P_{f_1} \cap P_{f_2}$, there is a stronger inequality of $P_{f_1*f_2}$. Hence, $P_{f_1*f_2} \subseteq P_{f_1} \cap P_{f_2}$.

$\mathbf{P_{f_1*f_2} \supseteq P_{f_1} \cap P_{f_2}}$

For any $\hat{x} \in P_{f_1} \cap P_{f_2}$ and any c_1, c_2 s.t. $c_1 + c_2 = \chi_X$, we must have $c_1^T\hat{x} \leqslant \max_{x\in P_{f_1}} c_1^T x$ and $c_2^T\hat{x} \leqslant \max_{x\in P_{f_2}} c_2^T x$. Hence, $\chi_X^T\hat{x} = c_1^T\hat{x} + c_2^T\hat{x} \leqslant (f_1)_{\mathrm{cup}}(c_1) + (f_2)_{\mathrm{cup}}(c_2)$. Thus every vector in $P_{f_1} \cap P_{f_2}$ belongs to $P_{f_1*f_2}$.

Next, for a given $X \subseteq S$, we will exhibit a vector $v \in P_{f_1} \cap P_{f_2}$ for which $\chi_X^T v = (f_1)_{\mathrm{cup}}(c_1) + (f_2)_{\mathrm{cup}}(c_2)$, for some c_1, c_2 s.t. $c_1 + c_2 = \chi_X$. This will prove that $f_1 * f_2$ is pt.

Let $\chi_X^T x$ reach its maximum among vectors of $P_{f_1} \cap P_{f_2}$ at some vertex v of $P_{f_1} \cap P_{f_2}$. The defining inequalities of $P_{f_1} \cap P_{f_2}$ are of those of P_{f_1} and P_{f_2} put together. So χ_X must be a nonnegative linear combination of the coefficient vectors of the supporting hyperplanes of v. Hence, $\chi_X = \sum \lambda_i \chi_{X_i} + \sum \mu_j \chi_{Y_j}, \lambda_i, \mu_j \geqslant 0$, where the first sum corresponds to supporting hyperplanes from P_{f_1} and the second from P_{f_2}. Let $c_1 \equiv \sum \lambda_i \chi_{X_i},\ c_2 \equiv \sum \mu_j \chi_{Y_j}$. Clearly $c_1^T v = \sum \lambda_i \chi_{X_i}^T v = \sum \lambda_i f_1(X_i)$ and $c_2^T v = \sum \mu_j \chi_{Y_j}^T v = \sum \mu_j f_2(X_i)$ and further these are the maximum values of $c_1^T x$ in P_{f_1} and of $c_2^T x$ in P_{f_2}. Thus $(f_1)_{\text{cup}}(c_1) + (f_2)_{\text{cup}}(c_2) = \chi_X^T v$ as required. $\square$

Remark. The above result can be more naturally seen using polarity ideas for polyhedral cones. We sketch the proof. It can be seen that $\mathscr{C}_{f_1 * f_2} = \mathscr{C}_{f_1} + \mathscr{C}_{f_2}$. So $\mathscr{C}^p_{f_1 * f_2} = \mathscr{C}^p_{f_1} \cap \mathscr{C}^p_{f_2}$. So $P_{f_1 * f_2} = P_{f_1} \cap P_{f_2}$. Use of Theorem 4.2 shows $f_1 * f_2$ is pt.

7. Conclusion

We have attempted to study a general class of discrete convex functions, namely the ones that can be extended to convex functionals, which we have called polyhedrally tight functions. Our basic approach is polyhedral. We use the generators of the associated dual cones of the concerned polyhedra through LDG structures. Some offshoots of this study are the recognition of the relative position of equivalent results such as Edmonds intersection theorem and Frank's Sandwich Theorem in this generalization. It turns out that the former is difficult to generalize unless we generalize the definition of convolution while the latter is routinely generalizable to all pt/dpt functions. We also show a function satisfies these theorems with all submodular pt/supermodular dpt functions respectively only if it is itself submodular. Using a strong 'hereditary regularity' property for LDG structures we are able to derive the integral version of Sandwich Theorem for submodular functions and the 1/2-integral version of the theorem for pseudomatroids.

References

[1] B. Bollobas, Linear Analysis, second ed., Cambridge University Press, London, 1999.
[2] J. Edmonds, Submodular functions, matroids, and certain polyhedra, in: R. Guy, H. Hanani, N. Sauer, J. Schönheim (Eds.), Proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications, Gordon and Breach, New York, 1970, pp. 69–87.
[3] A. Frank, An algorithm for submodular functions on graphs, Ann. Discrete Math. 16 (1982) 189–212.
[4] S. Fujishige, Submodular functions and optimization, Annals of Discrete Maths, North Holland, Amsterdam, New York, Oxford, Tokyo, 1991.
[5] L. Lovász, Submodular functions and convexity, in: A. Bachem, M. Grötschel, B. Korte (Eds.), Mathematical Programmming—The State of the Art, Springer, Berlin, 1983, pp. 235–257.

[6] R. Chandrasekharan, S.N. Kabadi, Pseudomatroids, Discrete Math. 71 (1988) 205–217.
[7] K. Murota, Discrete Convex Analysis, SIAM Monographs on Discrete Mathematics and Applications DT10, SIAM, Philadelphia, 2003.
[8] M. Sohoni, The shape of polyhedra in combinatorial optimization, Ph.D. thesis, Department of Computer Science and Engineering, Indian Institute of Technology, Bombay, 1992.